i , r ,

M L A G G V R SAGGCCAGTCTGTTTGCCTCCAACGCCATCTGACCCAGGTGAGCAAGAGG ATG CTG GCG GGG GGC GTG AGG S M P S P L L A C W Q P I L L V L G S 27 AGO ATG CCC AGO CCC CTC CTG GCC TGC TGG CAG CCC ATC CTC CTG CTG GTG CTG GGC TCA V L S G S A T G C P P R C E C S A Q D R 47 GTG CTG TCA GGC TCG GCC ACG GGC TGC CCC CCC CGC TGC GAG TGC TCC GCC CAG GAC CGC 271 A V L C H R K R F V A V P E G I P T E T 57 GCT GTG CTG TGC CAC CGC AAG CGC TTT GTG GCA GTC CCC GAG GGC ATC CCC ACC GAG ACG R L L D L G K N R I K T L N Q D E F A S 37 CSC CTG CTG GAC CTA GGC AAG AAC CSC ATC AAA ACG CTC AAC CAG GAC GAG TTC GCC AGC 391 FPHLEELELNENIVSAVEPG 107 TTC CCG CAC CTG GAG GAG CTG GAG CTC AAC GAG AAC ATC GTG AGC GCC GTG GAG CCC GGC 451 AFNNLFNLRTLGLRSNRLKL 127 GCC TTC AAC AAC CTC TTC AAC CTC CGG ACG CTG GGT CTC CGC AGC AAC CGC CTG AAG CTC 147 I P L G V F T G L S N L T K L D T R E N ATC CCG CTA GGC GTC TTC ACT GGC CTC AGC AAC CTG ACC AAG CTG GAC ACG AGG GAG AAC 167 K I V I L L D Y M F Q D L Y N L K S L E AAG ATC GTT ATC CTA CTG GAC TAC ATG TTT CAG GAC CTG TAC AAC CTC AAG TCA CTG GAG V G D N D L V Y I S H R A F S G L N S L 187 STT SGC GAC AAT GAC CTC STC TAC ATC TCT CAC CGC GCC TTC AGC GGC CTC AAC AGC CTG 207 EQLTLEKCNLTSIPTEALSH CAG CAG CTG ACT CTG GAG AAA TGC AAC CTG ACC TGC ATC CCC ACC GAG GCG CTG TCC CAC 227 LHGLIVLRLRHLNINAIRDY CTG CAC GGC CTC ATC GTC CTG AGG CTC CGG CAC CTC AAC ATC AAT GCC ATC CGG GAC TAC 311 247 S F K R L Y R L K V L E I S H W P Y L D 871 TCC TTC AAG AGG CTG TAC CGA CTC AAG GTC TTG GAG ATC TCC CAC TGG CCC TAC TTG GAC 267 T M T P N C L Y G L N L T S L S I T H C 931 ACC ATG ACA CCC AAC TGC CTC TAC GGC CTC AAC CTG ACG TCC CTG TCC ATC ACA CAC TGC 287 N L T A V P Y L A V R H L V Y L R F L N AAT CTG ACC GCT GTG CCC TAC CTG GCC GTC CGC CAC CTA GTC TAT CTC CGC TTC CTC AAC 991 307 LSYNPISTIEGSMLHELLRL CTC TCC TAC AAC CCC ATC AGC ACC ATT GAG GGC TCC ATG TTG CAT GAG CTG CTC CGG CTG 1051 327 Q E I Q L V G G Q L A V V E P Y A F R G CAG GAG ATC CAG CTG GTG GGC GGG CAG CTG GCC GTG GTG GAG CCC TAT GCC TTC GGC GGC 1111 347 L N Y L R V L N V S G N Q L T T L E E S CTC AAC TAC CTG CGC GTG CTC AAT GTC TCT GGC AAC CAG CTG ACC ACA CTG GAG GAA TCA 1171

TTGGGACCCAGCAGCAGCAGCAGTCAGGTGCATGCTGGGACCGGGACAGGGTGCGCCACCCCAGGCCCCA

V F H S V G N L E T L I L D S N P L .A C 367 STC TTC CAC TOS STS GGC AAC CTG GAG ACA CTC ATC CTG GAC TCC AAC CCS CTG GCC TGC 1231 DCRLLWYFRRRWRLNFNRQQ 387 PEFVQGKEFKDFPDV 407 ecc ace too occ ace coe sas titl site cas occ aas gas tite aas gad tite cot gat sig. 1351 L L P N Y F T C R R A R I R D R K A Q Q CTA CTG CCC AAC TAC TTC ACC TGC CGC CGC CGC ATC CGG GAC CGC AAG GCC CAG CAG .1411 V. F V D E G H T V Q F V C R A D G D P P 447 GTG TTT GTG GAC GAG GGC CAC ACG GTG CAG TTT GTG TGC CGG GCC GAT GGC GAC CCG CCG 1471 PAILWLSPRKHLVSAKSNGR 467 CCC GCC ATC CTC TGG CTC TCA CCC CGA AAG CAC CTG GTC TCA GCC AAG AGC AAT GGG CGG 1531 L T V F P D G T L E V R Y A Q V Q D N G 487 CTC ACA GTC TTC CCT GAT GGC ACG CTG GAG GTG CGC TAC GCC CAG GTA CAG GAC AAC GGC 1591 TYLCIAANAGGNDSMP 507 AGG TAC CTG TGC ATC GCG GCC AAC GCG GGC GGC AAC GAC TCC ATG CCC GCC CAC CTG CAT 1651 V R S Y S P D W P H Q P N K T F A F GTG CGC AGC TAC TCG CCC GAC TCG CCC CAT CAG CCC AAC AAG ACC TTC GCT TTC ATC TCC 1711 N Q P G E G E A N S T R A T V P F P F AAC CAG CCG GGC GAG GGA GAG GCC AAC AGC CGC GCC ACT GTG CCT TTC CCC TTC GAC 1771 I K T L I I A T T M G F I S F L G V V L 567 ATC AAG ACC CTC ATC ATC GCC ACC ACC ATG GGC TTC ATC TCT TTC CTG GGC GTC GTC CTC 1831 FCLVLLFLWSRGKGNTKHNI THE THE CHE CHE CHE THE CHE THE CHE THE AGE CHE GAG GGC AAC ACA AAG CAC AAC ATC 1891 EIEYVPRKSDAGISSADAPR GAG ATC GAG TAT GTG CCC CGA AAG TCG GAC GCA GGC ATC AGC TCC GCC GAC GCG CGC CGC 1951 615 KFNMKMI 1975 AAG TTC AAC ATG AAG ATG ATA TGA CTCACCACCTGCCTCCTTCTACCAGGACCTCAGAAGCCCAGACCTGGGGACCCCACCTACACAGGGGGCATTGACAGAC 2212 TGGAGTTTAAAGCCGACGACGACGCGGCAGAGTCAATAATTCAATAAAAAAGTTACGAACTTTCTCTGTAACTTG 2291

2351

Input file T81: Cutput File T811.pat

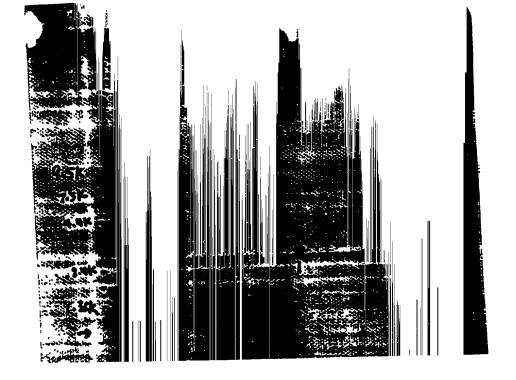
* Sequence length 979

	M A G S P																			
GAAT	SAATTCSGCACGAGGCCAGCCAGTCCSCCSGYMCGRRGCCCSGCTCSCTGGGGGCAGC ATG GCG GGG TCG CCG 7															72				
LLWGPRAGGGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC															25					
																				132
									R											45
Lin	CGG	CCG	CCC	CCC	GCG	CIC	TGC	GCG	CCG	CCG	GTA	AAG	GAG	CCC	CGC	GGC	CTA	AGC	CCA	192
A	s	P	P	L	А	Ξ	T	G	A	P	R	R	F	R	R	s	v	P	R	65
									GCT											2 52
_	-		,	~		**	^	=	L	,	2	۸	+	,	LF.	*	۲.	F	a	85
									cic											312
									A											105
GAA	gaa ogt cag gag ogg gog goc gag gog cag gag got gag gat cag cag gog cgc gtc 375 L a q L L R V W G A P R N S D P A L G L 125															3/2				
L	A	Q	L	L	R	A	W	G	A	5	2	И	s	ם	P	A	L	G	L	125
CIG	GCG	CAG	CIG	CIG	CCC	GTC	TGG	GGC	GCC	CCC	೧೧೮೦	AAC	TCI	gat	CCG	GCT	cic	GGC	TIG	432
_	_	_	_	_		_			_			_		,		-		-		145
									CYC G											492
a	P	A	A	L	A	A	Q	L	V	P	A	Þ	V	P	A	A	A	L	R	165
GAC	CCI	GCC	GCC	CTA	GCA	GCC	CAG	CIT	GIC	CCC	GCG	CCC	GIC	CCC	GCC	GCG	GCG	CIC	CGA	552
P	R	P	þ	v	Y	ם	D	G	P	A	G	P	D	A	E	E	A	G	D	185
									CCC											612
_	_	_	_	••	_	_	-		L	•	v	+		c	9	7	т.	a	G	205
																			GGA	672
									P										ם	225
AGC	GCG	GAC	TCC	: GAG	GGG	GIG	GCA	. GCC	CCG	CGC	CGC	cro	CGC	CGI	GCC	GCC	GAC	CAC	GAT	732
A	G	s	E	L	P	P	Ε	G	v	L	G	A	L	L	R	V	ĸ	R	L	245
																			CTA	792
_	_	_		_	_	••			-	-	,	,	_							261
									R CGC											840
CC	GCACTGCCCGGATGCCGTGCACCCTGGGACCCAGAAGTGCCCCGCCATCCCGGCACCAGGACTGCTCCCGGCAGCAC													91 9						
			~		~~~	19 ~~	***		CACC	~~~	~~	·~			-					979
216	فعلاسه	-	-11		المحار	المالك	باسلام		ساملاس.	-uni					~					2.5

= :	rcenc arm	4444	cy. as. the	
	T79	1 3	MLAGK LILACWQPILLLVLGSVLSGS .ATGCPPRCECSAQDR.	47
	D45913	1	MARLSTGKAAC.QVVLGLLITSLTESSILTSECPQLCVCEIRPWF	44
	T79	48	AVLCHRKRFVAVPEGIPTETRLLDLGKNRIKTLNQDEFAS	87
	D45913	45 '	TPQSTYREATTVDCNDLRLTRIPGNLSSDTQVLLLQSNNI	84
	T79	88	FPHLEELELNENIVSAVEPGAFNNLFNLRTLGLRSNRLKLIPLGVFTGLS	137
	D45913	85	AKTVDELQQLFNLTELDFSQNNFTNIKEVGLANLT	119
	т79	138	NLTKLDTRENKIVILLDYMFQDLYNLKSLEVGDNDLVYISHRAFSGLNSL	187
	D45913	120	QLTTLHLEENQISEMTDYCLQDLSNLQELYINHNQISTISANAFSGLKNL	169
	77 9	188	EQLTLEKCNLTSIPTEALSHLHGLIVLRLRHLNINAIRDYSFKRLYRLKV . :. : . . : : . : . : : . : : . : : : . : : . : : : : : : : : : : : : :	237
	54591 3	170	LRLHLNSNKLKVIDSRWFDSTPNLEILMIGENPVIGILDMNFRPLSNLRS	219
	17 79	238	LEISHWPYLDTMTPNCLYGLN.LTSLSITHCNLTAVPYLAVRHLVYLRFL	286
	045 913		LVLAG.MYLTDVPGNALVGLDSLESLSFYDNKLIKVPQLALQKVPNLKFL	
	79 -	287	NLSYNPISTIEGSMLHELLRLQEIQLVG.GQLAVVEFY:	323
	D 45913		DLNKNPIHKIQEGDFKNMLRLKELGINNMGELVSVDRYALDNLPELTKLE	
	፲279 ፲፬		AFRGLNYLRVLNVSGNQLTTLEESVFHSVGNLETLIL	
	1045913 ≟		ATNNPKLSYIHRLAFRSVPALESLMLNNNALNAVYQKTVESLPNLREISI	
	T79		DSNPLACDCRLLWVFRRRWRLNFNRQQPT.CATPEFVQGKEFKDFPDVLL . : :: : :	
	D45913		HSNPLRCDCVIHWINSNKTNIRFMEPLSMFCAMPPEYRGQQVKEVLI	
	T79		PNYFT.CRRARIRDRKAQQVFVDEGHTVQFVCRADGDPPPAILWLSPRKH .: . . : : : : : . : :	
	D45913		QDSSEQCLPMISHDTFPNHLNMDIGTTLFLDCRAMAEPEPEIYWVTPIGN LVSAKS.NGRLTVFPDGTLEVRYAQVQDNGTYLCIAANAGGNDSMPAHLH	
	T79			
	D45913		KITVETLSDKYKLSSEGTLEIANIQIEDSGRYTCVAQNVQGADTRVATIK V	515
	T79		:	
	D45913		VNGTLLDGAQVLKIYVKQTESHSILVSWKVNSNVMTSNLKWSSATMKIDN	
	T79	519	PNKTFAFISNQPGEGEANSTRA	540

The first state of the first sta

.....



77.3

T R P I L V I H D E O K G P E V T S N CC ACG CGT CCG ATC TTG GTC ATC CAC GAT GAA CAG AAG GGG CCG GAA GTG ACC TCC AAT A A L T L R N F C N W Q K Q H N P P S D GCT GCC CTC ACT CTG CGG AAC TTT TGC AAC TGG CAG AAG CAG CAC AAC CCA CCC AGT GAC 119 R D A E H Y D T A I L F T R Q D L C G S 59 CGG GAT GCA GAG CAC TAT GAC ACA GCA ATT CTT TTC ACC AGA CAG GAC TTG TGT GGG TCC 179 Q T C D T L G M A D V G T V C D P S R S 79 CAG ACA TOT GAT ACT CIT GGG ATG GCT GAT GTT GGA ACT GTG TGT GAT CCG AGC AGA AGC C S V I E D D G L Q A A F T T A H E L G 99 TGC TCC GTC ATA GAA GAT GAT GGT TTA CAA GCT GCC TTC ACC ACA GCC CAT GAA TTA GGC 299 H V F N M P H D D A K Q C A S L N G V N 119 CAC GIG TIT AAC ATG CCA CAT GAT GAT GCA AAG CAG TGT GCC AGC CTT AAT GGT GTG AAC 359 Q D S H M M A S M L S N L D H S Q P W S 139 CAG GAT TOO CAC ATG ATG GOG TOA ATG CTT TOO AAC CTG GAC CAC AGC CAG CCT TGG TOT PCSAYMITSFLDNGHGECLM 159 CCT TGC AGT GCC TAC ATG ATT ACA TCA TTT CTG GAT AAT GGT CAT GGG GAA TGT TTG ATG D K P Q N P I Q L P G D L P G T S Y D A 179 GAC AAG CCT CAG AAT CCC ATA CAG CTC CCA GGC GAT CTC CCT GGC ACC TCG TAC GAT GCC N R Q C Q F T F G E D S K H C P D A A S 199 AAC CGG CAG TGC CAG TTT ACA TTT GGG GAG GAC TCC AAA CAC TGC CCT GAT GCA GCC AGC TCSTLWCTGTSGGVLVCQTK 219 ACA TOT AGC ACC TTG TGG TGT ACC GGC ACC TCT GGT GGG GTG CTG GTG TGT CAA ACC AAA 239 H F P W A D G T S C G E G K W C I N G K CAC TTC CCG TGG GCG GAT GGC ACC AGC TGT GGA GAA GGG AAA TGG TGT ATC AAC GGC AAG C V N K T D R K H F D T P F H G S W G M TOT GTG AAC AAA ACC GAC AGA AAG CAT TTT GAT ACG CCT TTT CAT GGA AGC TGG GGA ATG WGPWGDCSRTCGGGVQYTMR TOG GGG CCT TGG GGA GAC TGT TCG AGA ACG TGC GGT GGA GGA GTC CAG TAC ACG ATG AGG ECDNPVPKNGGKYCEGKRVR 299 GAA TOT GAC AAC CCA GTC CCA AAG AAT GGA GGG AAG TAC TOT GAA GGC AAA CGA GTG CGC 899 319 Y R S C N L E D C P D N N G K T F R E E THE AGA TOO TGT AAC CTT GAG GAC TGT COA GAC AAT AAT GGA AAA ACC TTT AGA GAG GAA 959 Q C E A H N E F S K A S F G S G P A V E 339 CAA TOT GAA GCA CAC AAC GAG TIT TCA AAA GCT TCC TIT GGG AGT GGG CCT GCG GTG GAA 1019 WIPKYAGVSPKDRCKLICQA TIGG ATT CCC AAG TAC GCT GGC GTC TCA CCA AAG GAC AGG TGC AAG CTC ATC TGC CAA GCC 1079 K G I G Y F F V L Q P K V V D G T P C S 379 AAA GGC ATT GGC TAC TTC TTC GTT TTG CAG CCC AAG GTT GTA GAT GGT ACT CCA TGT AGC 1139

PDSTSVCVQGQCVKAGCDRI399 CCA GAT TOO ACC TOT GTC TOT GTG CAA GGA CAG TGT GTA AAA GCT GGT TGT GAT CGC ATC 1199 I D S K K K F D K C G V C G G N G S T C ATA GAC TOO AAA AAG AAG TYT GAT AAA TGT GGT GTT TGC GGG GGA AAT GGA TOT ACT TGT 1259 K K I S G S V T S A K P G Y H D I I T I 439 AAA AAA ATA TCA GGA TCA GTT ACT AGT GCA AAA CCT GGA TAT CAT GAT ATC ACA ATT 1319 PTGATNIEVKQRNQRGSRNN CCA ACT GGA GCC ACC AAC ATC GAA GTG AAA CAG CGG AAC CAG AGG GGA TCC AGG AAC AAT 1379 G S F L A I K A A D G T Y I L N G D Y T 479 GGC AGC TIT CIT GCC ATC AAA GCT GCT GAT GGC ACA TAT ATT CIT AAT GGT GAC TAC ACT 1439 L S T L E Q D I M Y K G V V L R Y S G S 499 TTG TCC ACC TTA GAG CAA GAC ATT ATG TAC AAA GGT GTT GTC TTG AGG TAC AGC GGC TCC 1499 SAALERIRSFSPLKEPLTIQ 519 TOT GOG GOA TTG GAA AGA ATT CGC AGO TTT AGO COT CTC AAA GAG CCC TTG ACC ATC CAG 1559 V L T V G N A L R P K I K Y T Y F V K K GTT CTT ACT GTG GGC AAT GCC CTT CGA CCT AAA ATT AAA TAC ACC TAC TTC GTA AAG AAG 1619 K K E S F N A I P T F S A W V I E E W G 559 AAG AAG GAA TOT TTO AAT GOT ATO COO ACT TTT TOA GOA TGG GTC ATT GAA GAG TGG GGC 1679 ECSKTCGKGYKKRSLKCLSH GAA TOT TOT AAG ACC TOT GGG AAG GGT TAC AAA AAA AGA AGC TTG AAG TGT CTG TCC CAT 1739 D G G V L S H E S C D P L K K P K H F I 599 GAT GGA GGG GTG TTA TCT CAT GAG AGC TGT GAT CCT TTA AAG AAA CCT AAA CAT TTC ATA 1799 609 D F C T M A E C S * 1829 GAC TIT TGC ACA ATG GCA GAA TGC AGT TAA GTGGTTTAAGTGGTGTTAGCTCTGAGGGCAAGGCAAAGTGAGGAAGGGCTGGTGCAGGGAAAGCAAGAAGGCTGGAGGG 1908 AAAAGTTAGAACTRTTACAACCCCTGTTTCCTGGTACTTATCAAATACTTAGTATCATGGGGTTGGGAAATGAAAAGT 2224 AGGAGAAAAGTGAGATTTTACTTAAGACCTGTTTTACTTTACCTCACTAACAATGGGGGGAGAAAGGAGTACAAATAGGA 2303 TCTTGACCAGCACTGTTTATGGCTGCTATGGTTTCAGAGAATGTTTATACATTATTTCTACCGAGAATTAAAACTTCA 2382 GATTGTTCAACATGAGAGAAAGGCTCAGCAACGTGAAATAACGCAAATGGCTTCCTCTTTCCTTTTTTTGGACCATCTCA 2461 GICTTATTIGIGIAATTCATTTIGAGGAAAAACAACTCCATGTATTTATTCAAGIGCATTAAAGTCTACAATGGAAA 2540 TACCATGTAACCCTGCTTTGGGAATATGGATGTAAAGAAGTAACTTGTGTCTCATGAAAATCAGTACAATCACAAAGG 2698

AGGATGAAAACGCCCGGAACAAAAATGACGTGTTAGAACAGGGTCCCACACGTTTTGGGACATTGAGATCACTTGTCTTG	2 777
TOGTOGGGAGGCTGCTGAGGGGGTAGCAGGTCCATCTCCAGCAGCTGGTCCAACAGTCGTATCCTGGTGAAATGTCTGTTC	2856
AGCTCTTCTGTGAGAATATGATTTTTTTCCATATGTATATAGTAAAATATGTTACTATAAATTACATGTACTTTATAAGT	2935
ATTGGTTTGGGTGTTCCTTCCAAGAAGGACTATAGTTAGT	3014
ATTTCTAATGAAAAAACTTTTAAATTATATCGCTTTTGTGGAAGTGCATATAAAATAGAGTATTTATACAATATATGT	30 93
TACTAGAAATAAAAGAACACTTTTTTTTAAAAAAAAAAA	3147

ANGO 71/ADAMIS-1 Comparison (90% Protein Sequence Identi	ty)
DOSMADFHGSGLKHYLLTLFSVAARFYKHPSIRNSISLVVVKILVIYEEQ 300	
TRPILVIHUEQ 11	
L KGPEVTSNAALTIRNFCSWOKOHNSPSDRDPEHYDTAILFTRODLCGSHT 350	
2 KGPEVTSNAALTIENFCNWQKQHNPPSDRDAEHYDTAILFTRQDLCGSQT 61	
1 CDTLCMADVGTVCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKH 400	
1 Caslagvsgdshimasmissldhsqpwspcsaymvtsfldnehgecimdk 450	
1 PONPIKLPSDLPGTLYDANROCOFTFGEESKHCPDAASTCTTLMCTGTSG 500	
. : :	
1 GLLVCQTXHFFWADGTSCGEGKWCVSGKCVNKTDMKHFATFVHGSWGFWG 550	
:	
1 PWGDCSRTCGGGVQYTMRECDNFVFKNGGKYCBGKRVRYRSCNTEDCPDN 600	
2 PWGDCSRTCGGGVQYTMRECDNFVPKNGGKYCEGKRVRYRSCNLEDCPDN 311	
1 NGKTFREEQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKURCKLTCEAKG 650	
1 IGYFFVLOPKVVDGTPCSPDSTSVCVQGQCVKAGCDRIIDSKKKFDKCGV 700	
01 CGGNGSTCKKMSGIVTSTRPGYHDIVTIPAGATNIEVKHRNORGSRNNGS 750	
il flairaadstytlmsmftlstleodltykstvlkyssssaalerirsfsp 800	
: :: :: :	
Ol lkeplitiqvimvghalrpkikftyfmkkktesfnaiptfsewvieemgec 850	1
	•
01 wspcsktcgkgykkrtlkovshdggvlsnesodplkkpkhyldfotltoc 950	١
	,
51 S* 951	
 1 08 S* 609	
40 S VV3	

FIG. 6

gtgc	ctac											Gly		tgt Cys		50
														ccc Pro		98
														gag Glu 45		146
														tgg Trp		194
	Gly													ttc Phe		242
tgg	gca Ala	gat Asp	ggc Gly	acc Thr	agc Ser	tgt Cys 85	gga Gly	gaa Glu	gly aaa	aag Lys	tgg Trp 90	tgt Cys	gtc Val	agt Ser	ggc Gly	290
	Cys													gtt Val		338
	Ser													acc Thr 125		386
400	ggt Gly													gtc Val		434
aag Lys	aac Asn	gga Gly 145	gly aaa	aag Lys	tac Tyr	tgt Cys	gaa Glu 150	Gly	aaa Lys	cga Arg	gtc Val	cgc Arg 155	tac Tyr	agg Arg	tcc Ser	482
		Ile					Asp					Thr		aga Arg		530
	Gln					Asn					Ala			gly		578
					Trp					Ala				cca Pro 205	Lys	626

	*															
gac Asp	agg Arg	tgc Cys	aag Lys 210	ctc Leu	acc Thr	tgt Cys	gaa Glu	gcc Ala 215	aaa Lys	ggc Gly	att Ile	ggc	tac Tyr 220	Phe	ttc Phe	674
gtc Val	tta Leu	cag Gln 225	ccc Pro	aag Lys	gtt Val	gta Val	gat Asp 230	ggc Gly	act Thr	ccc Pro	tgt Cys	agt Ser 235	cca Pro	gac Asp	tct Ser	722
acc Thr	tct Ser 240	gtc Val	tgt Cys	gtg Val	caa Gln	999 Gly 245	cag Gln	tgt Cys	gtg Val	aaa Lys	gct Ala 250	Gly	tgt Cys	gat Asp	cgc Arg	770
atc Ile 255	ata Ile	gac Asp	tcc Ser	aaa Lys	aag Lys 260	aag Lys	ttt Phe	gat Asp	aag Lys	tgt Cys 265	ggc	gtt Val	tgt Cys	gga Gly	gga Gly 270	818
aac Asn	ggt Gly	tcc Ser	aca Thr	tgc Cys 275	aag Lys	aag Lys	atg Met	tca Ser	gga Gly 280	ata Ile	gtc Val	act Thr	agt Ser	aca Thr 285	aga Arg	866
Pro C	Gly aaa	tat Tyr	cat His 290	gac Asp	att Ile	gtc Val	aca Thr	att Ile 295	cct Pro	gct Ala	gga Gly	gcc Ala	acc Thr 300	aac Asn	att Ile	914
Gln M	gtg Val	aaa Lys 305	His	cgg Arg	aat Asn	caa Gln	agg Arg 310	gly aaa	tcc Ser	aga Arg	aac Asn	aat Asn 315	ggc	agc Ser	ttt Phe	962
ctg Leu	gct Ala 320	att Ile	aga Arg	gcc Ala	gct Ala	gat Asp 325	Gly	acc Thr	tat Tyr	att Ile	ctg Leu 330	Asn	gga Gly	aac Asn	ttc Phe	1010
aČį	Leu	tcc Ser	aca Thr	cta Leu	gag Glu 340	Gln	gac Asp	ctc Leu	acc Thr	tac Tyr 345	Lys	ggt Gly	act Thr	gtc Val	tta Leu 350	1058
agg	tac Tyr	agt Ser	ggt Gly	tcc Ser 355	Ser	gct Ala	gcg Ala	g ctg Leu	gaa Glu 360	Arg	atc Ile	cgc Arg	agc Ser	ttt Phe 365	agt Ser	1106
cca Pro	cto Lev	: aaa Lys	a gaa Glu 370	ı Pro	tta Leu	acc Thr	ato	c cag e Glr 375	ı Val	ctt Leu	atg Met	gta : Val	ggc Gly 380	His	gct Ala	1154
			о Гуз					с Туг					Lys		gag Glu	1202
		e Ası					r Phe					l Ile			g tgg 1 Trp	1250

ggg gag tgc Gly Glu Cys 415	tcc aag Ser Lys	aca tg Thr Cy 420	c ggc s Gly	tca Ser	ggt Gly	tgg Trp 425	cag Gln	aga Arg	aga Arg	gta Val	5-5	1298
cag tgc aga Gln Cys Arg	gac att Asp Ile 439	e Asn Gl	a cac y His	cct Pro	gct Ala 440	tcc Ser	gaa Glu	tgt Cys	gca Ala	aag Lys 445	J	1346
gtg aag cca Val Lys Pro	gcc agt Ala Sei 450	acc ag Thr Ar	a cct g Pro	tgt Cys 455	gca Ala	gac Asp	ctt Leu	cct Pro	tgc Cys 460	cca Pro		1394
tgg cag gtg Trp Gln Val 465	. Gly Ası	t tgg to p Trp Se	a cca r Pro 470	tgt Cys	tcc Ser	aaa Lys	act Thr	tgc Cys 475	Gly	aag Lys	ggt Gly	1442
tac aag aag Tyr Lys Lys 480	g aga ac s Arg Th	c ttg aa r Leu Ly 48	rs Cys	gtg Val	tcc Ser	cac His	gat Asp 490	Gly	ggc	gtg Val	tta Leu	1490
tca aat gag Sea Asn Glu 495	g agc tg 1 Ser Cy	t gat co s Asp P: 500	t ttg	aag Lys	aag Lys	cca Pro 505	Lys	cat His	tac Tyr	att Ile	gac Asp 510	1538
till tgc aca Phiel Cys Th		r Gln C			gagg.	cgt	taga	ggac	aa g	gtag	cgtgg	1592
gg a ggggctg	atacact	.gag tgc	aagagt	a ct	ggag	ggat	cca	gtga	igtc	aaac	cagtaa	1652
gcagtgaggt	atagcaa	aga agt	atatat	a qo	ggat	acat	ago	aaag	gag	gtag	gatcagg	1712
açactaccct	accaatt	aca ttc	tgataa	g gt	agtt	aato	gago	caca	igta	gcat	ctgaaa	1772
qaccatacaq	agcacta	agg agc	cccaaa	ig ca	ictat	tagt	: ato	ctctt	ttc	ttat	atctat	1832
cgcccaaata	attttca	gag tct	ggcaga	a go	ccts	gttgo	act	gtac	ctaa	ctag	gatactt	1892
cttatcacaa	agattgg	gaa agg	caaago	a ga	aaaga	atggt	aaç	gacto	gggt	ttca	aacaag	1952
acttaatttc	aatcact	gga ggd	aaggag	g ag	gggga	acaaa	a caa	agato	catt	atto	gaagtc	2012
gétagttact	gtggttt	tac gga	aggttg	ga tç	gcato	catto	c cta	atcaa	acag	tgaa	aagttc	2072
aďČttqttca	acqtqac	aga aag	gctcat	c to	ccgts	gaaag	g ago	ctcct	gat	ttct	tcttac	2132
accatctcag	ttcttaa	acta tag	ttcate	gt to	gaggt	cagaa	a aca	aatto	catc	tati	cataaa	2132
atgtacattg	gaaaaaa	aaaa gtg	aagttt	a to	gaggt	cacao	c ata	aaaaa	actg	aagg	jaaacaa 	2232
tgagcaacat	gcctcct	get ttg	cttcct	C CT	gagg	gtaaa	a cci	zgeei	-999	gaci	gayyuu	2372
gtttaagatt	atccate	ggct cac	aagagg	ge ag	JLaac Staac	aalad	a car	rata	atca	gte	atctcac	2432
gaatggggta atgggaggct	tagagai	cag ggu	aggtag	ga ya	teet	ggaa	o att	aat.c	caac	agt	catatcc	2492
tggtgaatgt	getgeag	agge age	ctacto	ra o	agag	aatai	t ga	ctati	ttcc	ata	tgtatat	2552
gtatatagta	. cogcoc	atta cta	tgaati	ta di	atat	actt	t at	aaqt	attq	ata	tgtctgt	2612
tccttctaag	aaggaci	tata qtt	tataat	ta a	atqc	ctat	a at	aaca	tatt	tat	ttttata	2672
catttatttc	taatqa	taaa ac	tttaag	gt ta	atat	cgct	t tt	gtaa	aagt	gca	tataaaa	2732
atagagtatt	tataca:	atat atg	ttaact	ta g	aaat	aata	a aa	gaac	actt	ttg	aatgtgt	2792
atgcctattt	tctgga	gtgg gat	taact	tc t	gggc	aaga	a at	ctga	tgag	aca	caaacat	2852
tqqacttcaa	a gacagt	ttta aat	tttgg	gt a	aatg	aact	g ta	tttc	ctgt	tta	tagacgt	2912
actaataaaa	a aagaag	ttga tga	tgtct:	tt a	gtgg	taag	a tt	gtta	ctaa	tgt	ggttggc	2972
aaattqctqt	: aaaqaq	ccag ata	gtaag	ca t	ttat	ggca	t tg	tagg	ctat	ctt	tcctgcc	3032
acaaccatgt	gacagt	gagt gc	ttgta	gg a	ctga	gagc	a gc	cata	aatg	aca	tgtaaat	. 3092

ccagccctgc gagggcgcgc ggaccgggcg gaggtgttgt aggaggagac cgaggagggg gqctgggctg gggctggggc cgcgccggca agagagacat gcgattggtg accaagccga														120 180 240		
														ccc Pro 25		400
ctc Leu	ctg Leu	ctg Leu	gta Val 30	ctg Leu	ggc Gly	tca Ser	gtg Val	ctg Leu 35	tca Ser	ggc Gly	tct Ser	gct Ala	aca Thr 40	ggc Gly	tgc Cys	448
ccg Pro	ccc Pro	cgc Arg 45	tgc Cys	gag Glu	tgc Cys	tca Ser	gcg Ala 50	cag Gln	gac Asp	cga Arg	gcc Ala	gtg Val 55	ctc Leu	tgc Cys	cac His	496
	aaa Lys 60	cgc Arg	ttt Phe	gtg Val	gcg Ala	gtg Val 65	ccc Pro	gag Glu	ggc Gly	atc Ile	ccc Pro 70	acc Thr	gag Glu	act Thr	cgc Arg	544
ctg Leg 75	ctg Leu	gac Asp	ctg Leu	ggc Gly	aaa Lys 80	aac Asn	cgc Arg	atc Ile	aag Lys	aca Thr 85	ctc Leu	aac Asn	cag Gln	gac Asp	gag Glu 90	592
														aac Asn 105		640
gtg Val														ctg Leu		688
act Thr	ctg Leu	999 Gly 125	ctg Leu	cgc Arg	agc Ser	aac Asn	cgc Arg 130	ctg Leu	aag Lys	ctt Leu	atc Ile	ccg Pro 135	ctg Leu	ggc	gtc Val	736
ttc Phe	acc Thr 140	ggc Gly	ctc Leu	agc Ser	aac Asn	ttg Leu 145	acc Thr	aag Lys	ctg Leu	gac Asp	atc Ile 150	agt Ser	gag Glu	aac Asn	aag Lys	784
	Val										Leu			ctc Leu		832
					Asp					Tyr				cga Arg 185	Ala	880

	•															
ttc Phe	agc Ser	ggc Gly	ctc Leu 190	aac Asn	agc Ser	ctg Leu	gaa Glu	cag Gln 195	ctg Leu	acg Thr	ctg Leu	gag Glu	aaa Lys 200	tgc Cys	aat Asn	928
ctg Leu	acc Thr	tcc Ser	atc	ccc Pro	acg Thr	gag Glu	Ala	ctc	tcc Ser	cac His	ctg Leu	His	ggc Gly	ctc Leu	atc Ile	976
atc	cta	205 caa	cta	cga	cat	ctc	210 aac	atc	aat	gcc	atc	215 agg	gac	tac	tcc	1024
Val	Leu 220	Arg	Leu	Arg	His	Leu 225	Asn	Ile	Asn	Āla	Ile 230	Arg	Asp	Tyr	Ser	
ttc Phe 235	aag Lys	agg Arg	ctg Leu	tac Tyr	cga Arg 240	ctt Leu	aag Lys	gtc Val	tta Leu	gag Glu 245	atc Ile	tcc Ser	cac His	tgg Trp	ccc Pro 250	1072
tac Tyr	ctg Leu	gac Asp	acc Thr	ata Ile 255	acc Thr	ccc Pro	cgg Arg	acg Thr	cgt Arg 260	Gly	tcg Ser	ac				1110
ring archae.																

The start given start starts start and start start

ctcctggatg tgcgcagccg cagagcgctg ctgctgtgcc taatacccat cgctgcgcac 6 ttgacagcca gtccgcccgt ccggagcccg gctcgttggg gcagc atg gcg ggg tcg 1 Met Ala Gly Ser 1														60 117		
ccg Pro 5	ctg Leu	ctc Leu	tgc Cys	Gly 333	ccg Pro 10	cgg Arg	gcc Ala	Gly aaa	ggc Gly	gtc Val 15	ggc	att Ile	ttg Leu	gtg Val	ctg Leu 20	165
ctg Leu	ctc Leu	ttg Leu	ggc Gly	ctt Leu 25	ctg Leu	agg Arg	ctg Leu	ccc Pro	ccc Pro 30	acc Thr	ctg Leu	tca Ser	gcg Ala	agg Arg 35	ccc Pro	213
gtg Val	aag Lys	gag Glu	ccc Pro 40	cgc Arg	agt Ser	ctg Leu	agc Ser	gca Ala 45	gca Ala	tcc Ser	gcg Ala	ccc Pro	ttg Leu 50	gtt Val	gag Glu	261
acg Thr	agc Ser	act Thr 55	ccc Pro	ctc Leu	cgc Arg	ttg Leu	cgt Arg 60	cgg Arg	gcc Ala	gtg Val	ccc Pro	cga Arg 65	gga Gly	gag Glu	gcg Ala	309
geg Ala	ggt Gly 70	gcg Ala	gtg Val	cag Gln	gag Glu	ctg Leu 75	gcg Ala	cgg Arg	gcg Ala	ctg Leu	gcg Ala 80	cac His	ctg Leu	ctg Leu	gag Glu	357
9 9 9 A I ā 85	gag Glu	aga Arg	cag Gln	gaa Glu	cgc Arg 90	gcg Ala	cgt Arg	gct Ala	gag Glu	gcg Ala 95	cag Gln	gag Glu	gct Ala	gag Glu	gat Asp 100	405
cag Gla	cag Gln	gcg Ala	cgt Arg	gtc Val 105	ctg Leu	gcg Ala	cag Gln	ctg Leu	ctg Leu 110	cgc Arg	gcc Ala	tgg Trp	ggc Gly	tct Ser 115	ccg Pro	453
cği	gcc Ala	tcg Ser	gac Asp 120	ccg Pro	ccc Pro	ttg Leu	gcc Ala	ccc Pro 125	gac Asp	gat Asp	gac Asp	ccg Pro	gac Asp 130	gct Ala	cca Pro	501
gct Ala	gca Ala	cag Gln 135	Leu	gcc Ala	cgt Arg	gct Ala	ctg Leu 140	Leu	cga Arg	gct Ala	cgc Arg	cta Leu 145	Asp	ccc Pro	ggc	549
ccc	cag Glr 150	Cys	atg Met	atg Met	atg Met	gcc Ala 155	Pro	ctg Leu	gcc Ala	caç Glr	g acc n Thr 160	Ser	agg Arg	atg Met	ccg Pro	597
gcg Ala 165	a Thr	aga Arg	ı cto g Lei	c ctg 1 Leu	acg Thr	Trp	g acc	ctg Leu	ı ago	tgo Cys 175	3	iggta	ictt	gcta	raaaca	g 650
gto	ggaco	cagg	attt	gggt	cc c	gagg	gtgco	cc cc	tgag	gaac	g tad	ctggg	gggc	tate	cegate getacg ceetg	c 770